REMARKS

This amendment is in response to the Examiner's Office Action dated 4/22/2004. This amendment should obviate outstanding issues and make the claims allowable. Reconsideration of this application is respectfully requested in view of the foregoing amendment and the remarks that follow.

STATUS OF CLAIMS

Claims 1-25 are pending.

Claims 1-25 stand rejected under 35 U.S.C. § 102(e) as being anticipated by Kraft et al. (USP 6,516,312).

OVERVIEW OF CLAIMED INVENTION

The presently claimed invention allows a web crawler to accurately mimic real users, by relying on past user accesses to the Web sites to be crawled. This approach results in a web crawler capable of automatically accessing all the content that a real user would have access to.

In one embodiment, the present invention enumerates parameter combinations for automated access to World Wide Web content that a real user access would have accessed. Parameters combinations are based on input values that the real user has provided to input fields of a World Wide Web site. Parameter combinations are selected in a manner such that automated access patterns are "equivalent" real user access patterns. A log file maintains at least one set of parameters corresponding to a specific instance of real user interactions with a World Wide Web site. This log file is then analyzed to enumerate possible parameter combinations for

achieving automated access to the World Wide Web content "semantically equivalent" in nature to real user access patterns.

In another embodiment, a method of determining entries for input to an HTML form in pursuit of automated accesses to content contained in a Web database, is provided. Real user entries provided to an HTML form are logged and analyzed to enumerate combinations of entries for automatically populating an HTML file and the subsequent automated accesses to web content.

<u>REJECTIONS UNDER 35 U.S.C. § 102(e)</u>

The examiner has rejected claims 1-25 under 35 U.S.C. § 102(e) as being anticipated by Kraft et al. (USP 6,516,312). To be properly rejected under 35 U.S.C. §102, each and every element of claims must be disclosed in a single cited reference. The applicants, however, contend that the presently claimed invention cannot be anticipated in view of the '312 reference. The Kraft et al reference (hereafter Kraft) is primarily cited for its provision of new search queries generated from a domain-specific user query that was previously, dynamically associated with keywords. The Kraft reference teaches away from the present invention by generating a new search result from a set of previously prepared abstracts and by providing additional, supplemental information to each user query. Since a search engine repository is updated with this additional information, subsequent executions of the same user query will not, and are not intended to, generate the same or equivalent search results. In direct opposition, the present invention provides a method for consistently accessing the same instance of web content.

The Kraft reference is also cited for its provision of an "optimal" query string generated from a domain-specific query through the steps of: composing a complex, boolean query string of keywords, executing it against the World Wide Web (WWW), and calibrating the quantity of search results to a "manageable" size. A search engine repository stores these calibrated search

queries by incorporating them along with keywords from the domain-specific query into a search abstract. Subsequently disclosed are links between new queries and both domain-specific keywords and query strings that were incorporated in these search abstracts.

In regards to independent claims 1, 7, 8, and 9, the examiner has cited figure 3 of Kraft for its inclusion of a log file component containing an abstract component obtained by a web crawler and a browser component. Discussion accompanying figure 3, as cited by the examiner in column 7, lines 4-6 and column 7, lines 54-60, discloses the use of metadata, defined as Uniform Resource Locators (URLs) and keywords by Kraft, from previously encountered web pages to create abstracts contained in the log file component. The Kraft reference teaches an incremental process of extracting this metadata from previous web page hits and incorporating in an abstract, selectively-joined portions of extracted metadata.

By contrast, the present invention discloses an ordered set of parameters in a log file chosen such that automated access to the same WWW content as would be accessed manually, by a real user, is provided. Each parameter stored in a log file of the present invention is comprised of a name and associated value, specifically, an input field name in a WWW form and an associated input value to this field. In other words, the presently claimed invention seeks to reverse engineer a manual access of web content by automatically answering a question (i.e. input field name) presented by a web site with an answer (i.e. input value) that is based on a stored set of user responses (i.e. parameters values) to the same question (i.e. parameter name) presented by the same WWW form. A combination, as specified by the present invention, is *a set of parameters that are individually input to a web form*, whereas the combination disclosed by Kraft is number of distinct URLs and keywords *combined to create a single query string*.

Furthermore, Kraft teaches the automatic provision of a new search result to a browser for execution or to a web crawler for traversal. Therefore, applicants contend that abstracts

contained in the log file maintained in search service provider cited by the examiner cannot be used to determine parameter combinations, nor can they be used in attaining access to web content, wherein access is automated or otherwise.

In regard to independent claim 5, figure 3 of the Kraft reference is also cited by the examiner as suggesting a proxy server used in the description of the present invention. A proxy server of the present invention refers to a computer or program that is transparent to a client; a client does not see or know that a proxy server exists. Instead, a client sees web content produced by a web server to which the client is connected. A proxy server records communication between a web server and client silently and transparently (i.e. without requiring a client to know of its existence). In contrast, the search service provider pointed to by the examiner in the Kraft reference is, in fact, a web server directly providing content in response to a client's request. It is implied, therefore, that a client is aware of the search service provider's existence by the fact that a client issues a request for content directly to the search service provider. Thus, a search service provider cannot be a proxy server for the following reasons: it provides content; a client is aware of a direct connection to it; and it does not intercept communications, transparently or otherwise.

With respect to independent claims 2, 3, and 6, the examiner has cited figure 6a as illustrating parameters. Specifically, the examiner has equated parameters disclosed in the present invention with a text string and arbitrary URLs containing the same text string. Such a text string, for example, "RMI" in figure 6a as cited by the examiner, is not a parameter but rather, a simple keyword. A parameter of the present invention requires, as previously discussed, both a name (e.g. "zip code") and an associated value (e.g. "95120") appropriate for the name. In order for a web crawler to gain automated access to certain web content, the present invention teaches the determination of a value for a name component of a given parameter requested by a

web site, for example a value for a zip code. In essence, when a Web site presents the question "What is the zip code?", a crawler reads previous responses stored for this question, answer the question with a value or values, "95120", based on what it read.

The examiner has pointed to phrases such as "combinations of entries" and "the combination of keywords with URLs" suggesting equivalence to a combination of parameters. However, a combination of parameters as disclosed by the present invention makes a provision for a plurality of distinct pairs of input field name and corresponding input field value. Since multiple input fields are to be filled out, a combination or a set of multiple input values is required. For example, if a form has the input fields "zip code" and "city name", appropriate combinations might be "95120" and "San Jose" as well as "94109" and "San Francisco." Because keywords and URLs are not parameters (i.e. they are not input fields with appropriately specified input values), their combination cannot be used to appropriately fill out forms having fields requiring input. Additionally, keywords cited in figure 6a of the Kraft reference are not parameters because there are no input fields requiring input values.

With regards to independent claims 4, 10, and 11, the examiner has cited figure 6a of Kraft as suggesting both limit and unlimited text entries. A text entry of the present invention is limited if a vocabulary or range associated with an input field is restricted to a given set of words or input values, respectively; and unlimited otherwise. For instance, an input field "zip code" may have restricted input values "95120", "95121, and "95122" and therefore the input field can only accept a limited text entry. In contrast, an input field "author's name" may be unlimited in vocabulary, as it is impossible to enumerate all possible author names. A key difference of Kraft lies the fact that abstracts and articles to which they are linked are neither limited only by a number of words, nor are they limited by a vocabulary associated with words; any number of words from a vocabulary of any language can be a part of an abstract or article. Furthermore, an

abstract is not an input value to a specific input field on a WWW form and therefore, cannot be an entry to such a form.

The previous arguments presented with respect to independent claims 1, 7, 8, and 9, substantially apply to dependent claims 12, 17, and 20. Additionally, arguments presented previously for independent claim 5 are substantially applicable as arguments against the rejection of dependent claims 16, 19, and 24. Similarly, arguments made above for independent claims 2, 3, and 6 and arguments made for independent claims 4, 10, and 11 substantially apply to dependent claims 13, 14, 17, 21, 22, and 25 and to dependent claims 15, 18, and 23, respectively.
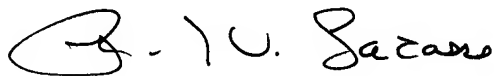
SUMMARY

As has been detailed above, none of the references, cited or applied, provide for the specific claimed details of applicants' presently claimed invention, nor renders them obvious. It is believed that this case is in condition for allowance and reconsideration thereof and early issuance is respectfully requested.

As this amendment has been timely filed within the set period of response, no petition for extension of time or associated fee is required. However, the Commissioner is hereby authorized to charge any deficiencies in the fees provided to Deposit Account No. 09-0441.

If it is felt that an interview would expedite prosecution of this application, please do not hesitate to contact applicants' representative at the below number.

Respectfully submitted,

Randy W. Lacasse
Registration No. 34,368

1725 Duke Street
Suite 650
Alexandria, Virginia 22314
(703) 838-7683
June 15, 2004